Functional Magnetic Resonance Imaging of Human Auditory Cortex

J. R. Binder, MD,* S. M. Rao, PhD,* T. A. Hammeke, PhD,* F. Z. Yetkin, MD,† A. Jesmanowicz, PhD,‡
P. A. Bandettini, BS,‡ E. C. Wong, PhD,‡ L. D. Estkowski, BS,† M. D. Goldstein, PhD,§
V. M. Haughton, MD,† and J. S. Hyde, PhD‡

Magnetic resonance imaging methods recently demonstrated regional cerebral signal changes in response to limb movement and visual stimulation, attributed to blood flow enhancement. We studied 5 normal subjects scanned while listening to auditory stimuli including nonspeech noise, meaningless speech sounds, single words, and narrative text. Imaged regions included the lateral aspects of both hemispheres. Signal changes in the superior temporal gyrus and superior temporal sulcus were observed bilaterally in all subjects. Speech stimuli were associated with significantly more widespread signal changes than was the noise stimulus, while no consistent differences were observed between responses to different speech stimuli. Considerable intersubject variability in the topography of signal changes was observed. These observations confirm the utility of magnetic resonance imaging in the study of human brain structure-function relationships and emphasize the role of the superior temporal gyrus in perception of acoustic-phonetic features of speech, rather than processing of semantic features.

Binder JR, Rao SM, Hammeke TA, Yetkin FZ, Jesmanowicz A, Bandettini PA, Wong EC, Estkowski LD, Goldstein MD, Haughton VM, Hyde JS. Functional magnetic resonance imaging of human auditory cortex. Ann Neurol 1994;35:662–672

The anatomy and physiology underlying speech perception have been studied with a variety of methods, each providing important clues about the organization of this complex function. Human studies have included cytoarchitectural mapping of auditory cortex in cadaver brains [1, 2], behavioral-anatomical correlations in brain-injured persons [3–6], evoked potentials to auditory stimuli [7–9], and mapping of regional blood flow and metabolism changes during audition [10–14]. Structural studies in animals have begun to unravel the complex hierarchical connections among auditory cortical areas [15–17], and cell recordings in awake animals have provided detailed data regarding neuronal events in response to speech and speech-related stimuli [18–21]. While incorporating fundamentally different types of test paradigms and measurements, many of these studies suggested a major role for perisylvian temporal cortex in the analysis of auditory speech stimuli [4, 9, 14, 22]. Both right and left temporal regions appear to participate to some degree [9, 11, 13, 14]. Processing of complex stimuli such as consonant-vowel clusters or words appears to require the activity of more widespread brain areas than does processing of certain non-speech sounds [12, 22].

This report presents preliminary findings using magnetic resonance imaging techniques to visualize brain regions involved in auditory speech perception. The method, herein referred to as "functional magnetic resonance imaging" (FMRI), measures regional increases in magnetic resonance signal during brain activity. In explaining these regional changes, previous investigators suggested that increased blood flow and oxygenation at the capillary venous level in metabolically active tissue [23] leads to decreased intravoxel dephasing, resulting in increased signal [24–28]. FMRI has previously been used to image brain activity during finger movements [25, 29] and during primary visual stimulation [26–28]. It is unknown whether FMRI can detect activity in sensory association cortices, particularly since reported blood flow changes in such areas have been of much smaller magnitude than those occurring in primary sensory and motor areas [13].

With its capability for producing high-resolution images, FMRI could contribute to an understanding of the detailed functional anatomy of auditory processing. As an initial investigation of auditory cortex, we measured magnetic resonance signal changes in the lateral aspect of both temporal lobes as subjects passively listened to speech and nonspeech stimuli. Stimuli differed in both semantic content and acoustic feature (frequency modulation) content. The study was designed to address the following questions: (1) Can

From the Departments of *Neurology, †Radiology, ‡Biophysics, and §Psychiatry, Medical College of Wisconsin, Milwaukee, WI.

Received Mar 15, 1993, and in revised form Sep 3. Accepted for publication Nov 4, 1993.

Address correspondence to Dr Binder, Department of Neurology, Medical College of Wisconsin, 9200 W. Wisconsin Avenue, Milwaukee, WI 53226.

FMRI demonstrate activity in primary and association auditory cortex? (2) If so, what are some salient stimulus variables determining activation in these regions? (3) Are there observable differences between individuals in the location and extent of activity measured by FMRI?

Materials and Methods

Subjects

Subjects were 5 healthy individuals (3 women, 2 men), aged 25 to 30 years, who had no history of neurological or auditory symptoms. Screening neurological examinations revealed normal findings. Individual data, including Edinburgh Handedness Inventory laterality quotients [30], are provided in the Table. Subjects were recruited on a voluntary basis, gave written informed consent according to institutional guidelines, and were paid a small hourly stipend. All studies received prior approval by the institutional human research review committee.

Apparatus and Scanning Procedures

Imaging was performed on a General Electric 1.5-T Signa scanner using a 30-cm three-axis head gradient coil designed for rapid gradient switching [31], and an elliptical endcapped quadrature radiofrequency (rf) coil designed for whole-volume brain imaging [32]. A blipped gradient-echo echoplanar sequence was used for functional imaging, allowing complete acquisition of a 64 × 64 pixel image in 40 msec. The field of view (FOV) was 24 cm and slice thickness 10 mm, yielding voxel dimensions of 3.75 × 3.75 × 10 mm. Symmetrical lateral sagittal slices of the left and right hemispheres were obtained, centered at positions 8 mm medial to the most lateral point of the temporal lobe on each side. Sixty-four sequential images of each slice were collected, alternating left- and right-sided acquisitions, with an interscan temporal spacing, or repetition time (TR), of 3 seconds for each side.

Subjects lay in the scanner with eyes closed and room lights extinguished. Digitally recorded auditory stimuli were played at precise intervals using a Macintosh microcomputer, amplified, and delivered to the subject via air conduction through a semirigid 1-cm bore plastic tube. The tube conducted the sound stimulus approximately 20 ft from the control room wall to the subject, at which point a Y-connector split the tube for binaural stimulation through a tightly fitting headset with occlusive earplugs to further reduce scanner noise exposure. A 10-band frequency filter was used, with uniform settings across subjects, to replace high-frequency loss in the auditory stimulus due to the tube conduction system.

Subject Data

Subject	Sex	Age (yr)	LQ	Test Order
1	F	25	78	N-P-W
2	F	28	90	N-W
3	M	30	65	N-P-W
4	M	30	100	N-W-P-T
5	F	26	78	N-W-P-T

LQ = laterality quotient [30]; N = noise; P = pseudowords; W = words; T = text.

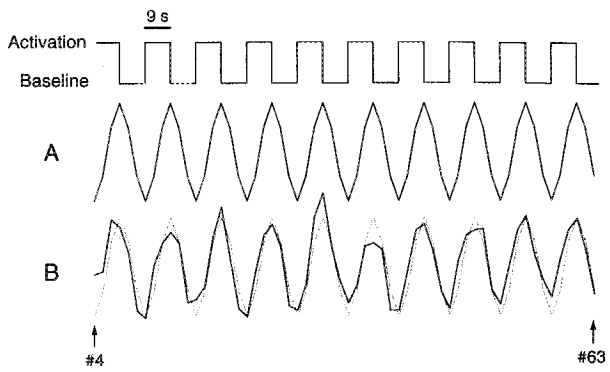


Fig 1. Timing of stimulus presentations and data analysis. Beginning with image 4, 9-second periods of stimulus presentation ("activation") alternated with 9-second periods of baseline for a total ten cycles. (A) Reference sinusoidal model of signal change over time. Note the alignment of peaks with ends of activation periods and alignment of troughs with ends of baseline periods. (B) Experimental data from one voxel in the left superior temporal gyrus of Subject 3, superimposed on the reference function (r = 0.956).

Scanning began with acquisition of standard 256 × 128 pixel, T1-weighted, 5-mm GRASS (gradient-recalled at steady state) images to be used for later anatomical localization, located at the same plane and center position as the echoplanar images. Prior to echoplanar imaging, subjects were informed that they would hear various sounds through the earphones, and that they should "listen carefully to the sounds in the earphones." Each 64-image echoplanar series consisted of multiple periods of "baseline," during which subjects heard only the ambient scanner noise, alternating with periods of "activation," during which one of four types of prepared auditory stimulus was delivered (Fig 1). Each series began with 3 baseline images (9-second interval) allowing magnetic resonance signal equilibrium to be reached, followed by 60 images during which activation alternated with baseline every 9 seconds (6 images/cycle, 18 sec/cycle, ten cycles). Background scanner noise was constant throughout all baseline and activation periods. Although the onset of the first activation period was manually initiated (and therefore somewhat variable relative to the timing of image acquisition), subsequent baseline and activation periods were digitally triggered and therefore precisely regular. The final image in the series was a reference image containing a selective excitation grid enabling estimation of any gross anatomical distortion resulting from the echoplanar technique [33].

Stimuli

Stimulus amplitude, which remained constant across subjects, averaged 117 dB sound pressure level (SPL) at the distal end of the audio system for all four types of stimuli. Peak amplitude of the background scanner noise (which remained constant through all baseline and activation periods) was approximately 95 dB SPL. The occlusive earplugs attenuated both scanner noise and test stimuli by an estimated 25 dB SPL, resulting in approximate tympanic SPLs of 97 dB for the

experimental stimuli and 70 dB for the background noise. Subjects reported perceiving the activation stimuli as subjectively louder than the background scanner noise. One of the following four stimulus types was delivered during each echoplanar series:

1. *Words.* These were 90 monosyllabic, concrete English nouns of medium frequency (mean = 40 occurrences/million; range, 10–88 [34]), for example, *barn, shore, box.* The digitial recordings featured a male voice, and were each edited to an approximate duration of 0.8 second. Playback during the activation periods was at a rate of 1 word/sec, allowing presentation of 9 words during each activation period.
2. *Pseudowords.* These were 90 monosyllabic nonsense phoneme strings, derived by rearrangement of the constituent phonemes of the word stimuli, for example, *narb, orsh, skob.* The voice used, average SPL, duration, and stimulus rate were identical to those for the word stimuli.
3. *Text.* This consisted of a 257-word narrative story adapted from a work of Tolstoy and containing many concrete nouns and highly imageable noun phrases and action descriptions (see Appendix). The story was divided into ten 9-second sentences for presentation during the activation periods of one image series. Individual voice and SPL characteristics matched those of the other stimuli. Constituent words were presented at a mean rate of 3.6 syllables/sec.
4. *Noise.* This nonspeech stimulus consisted of "white noise" matched in SPL, duration, spectral range (50–5,000 Hz), and stimulus rate to the word and pseudoword stimuli, but lacking any frequency modulation features.

The order of test stimuli for each subject is provided in the Table. The order of word and pseudoword presentation was varied across subjects to minimize order effects on this comparison. Several subjects were tested twice with each stimulus, although no stimuli were repeated until all had been presented. All data in this article describe responses to initial presentations.

Data Analysis

Previous FMRI observations suggested a rise latency in sensory cortex of 8 to 9 seconds from stimulus onset to maximal signal change, and a fall latency of 9 to 10 seconds from stimulus cessation to baseline signal [35]. Signal changes over time during each 18-second cycle in the present study were therefore predicted to resemble a sinusoidal function; pilot studies confirmed this prediction (see Fig 1).

FMRI studies have generally produced a "functional image" by computing for each voxel the difference in mean signal between single baseline and activation periods [26–28, 36]. An alternative method, developed in our laboratory and used in the present study, correlates the periodic data obtained from multiple alternating periods of baseline and activation with periodic reference functions (sine functions) to identify voxels responding to the stimulus [29, 37]. This method improves the signal-to-noise ratio by incorporating a larger set of data points per analysis (i.e., all data points acquired). Like the voxel data, the reference functions used in this study consisted of ten cycles defined by six data points/cycle and were phase-adjusted so that trough points were temporally aligned with the last image of the baseline (no stimulus) periods (see Fig 1). Because the precise phase of the stimulus cycle relative to image acquisitions was unknown, and probably varied from series to series by as much as 1 second, ten reference sinusoidal functions were used and were phase-shifted over a range of ± 0.5 second, in increments of 100 msec, around the expected phase. The best r generated for each voxel was used for further analysis, regardless of which reference function produced the correlation. This method compensates for imprecision in the manual triggering of the stimulus cycles and also reduces the requirement that all brain regions respond with precisely the same temporal phase shift relative to the stimulus [38].

Only those voxels with $|r| \geq 0.50$ were analyzed further. This cutoff corresponds to $p \leq 0.0001$ for a 60-image series [37], a value selected to reduce the likelihood of type I error due to the large number of voxels (approximately 500) considered in each analysis. Images representing the magnitude of stimulus-locked signal change in these voxels were constructed by setting the brightness of the corresponding pixels to $K(r\sigma)$, where σ is the standard deviation of the voxel data (proportional to the magnitude of signal change), and K is a scale constant [37]. For anatomical localization, the brightness scale of the functional images was converted to a polychromatic scale, and functional images were superimposed on T1-weighted GRASS images of the same brain slices after interpolating all images to 256 × 256 matrices. Inspection of the selective excitation grid images showed no appreciable anatomical distortion in the regions containing superior temporal, middle temporal, and frontal lobes; consequently registration of the functional and anatomical images was considered to be very good.

Results

General Observations

Signal increase that followed the onset of sound stimuli (hereafter referred to as "activation") was consistently observed in the superior temporal gyrus (STG) of all subjects, and ranged from approximately 1 to 5% of baseline signal values (Fig 2). Correlations between reference sinusoidal functions and actual signal data frequently exceeded 0.90 for those voxels with strong responses (see Fig 1), suggesting that sinusoidal functions are a good approximation of the physiological response under these activation conditions. Figure 3A illustrates a "baseline image" derived by averaging 10 echoplanar images taken during baseline periods, and Figure 3B an "activation image" derived by averaging 10 images taken during peak activation by speech stimuli. The result of a simple subtraction between these two images is shown in Figure 3C, which demonstrates the main focus of signal change in the STG despite considerable residual "noise" in the image. Figure 3D shows the improved functional image obtained using the cross-correlation thresholding technique described above, which more clearly delineates areas of stimulus-related signal fluctuation.

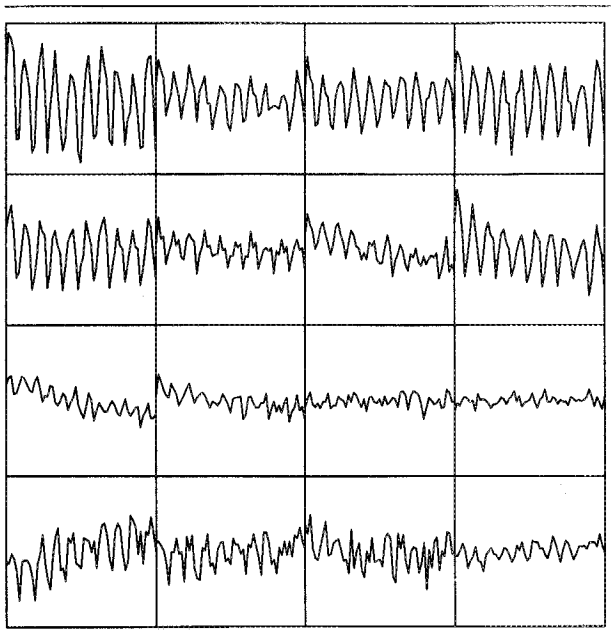


Fig 2. Raw data showing percent signal change (vertical axes) over time (horizontal axes) in 16 contiguous voxels of the left superior temporal gyrus of Subject 5 during pseudoword presentation (see Fig 4). Vertical axis range for all boxes equals 10%.

Regions of activation for each subject are illustrated in Figure 4, using a red-yellow scale to indicate positively correlated signal activity and a blue-cyan scale to indicate negatively correlated (signal decrement during stimulation) activity. The great majority of the temporally correlated changes were positive in relation to the stimulus and were centered over cortex rather than white matter. Signal changes were often particularly strong over sulci, where the cortical ribbon was more likely to traverse the entire thickness of the slice, minimizing volume averaging effects.

Stimulus Effects

The temporal lobe area (number of voxels) activated by either word or pseudoword stimuli was significantly greater than the area activated by the white noise stimulus, as shown by two-way repeated measures analysis of variance (ANOVA) ($F(2,6) = 37.7$, $p < 0.0001$) followed by t tests (Fig 5). With white noise stimulation, signal increases were characteristically confined to the dorsal aspect of the STG, in most cases centered on the transverse temporal (Heschl's) gyrus (TTG) (see Fig 4). Noise activation was somewhat more widespread in Subjects 1 and 3, including small areas ante-

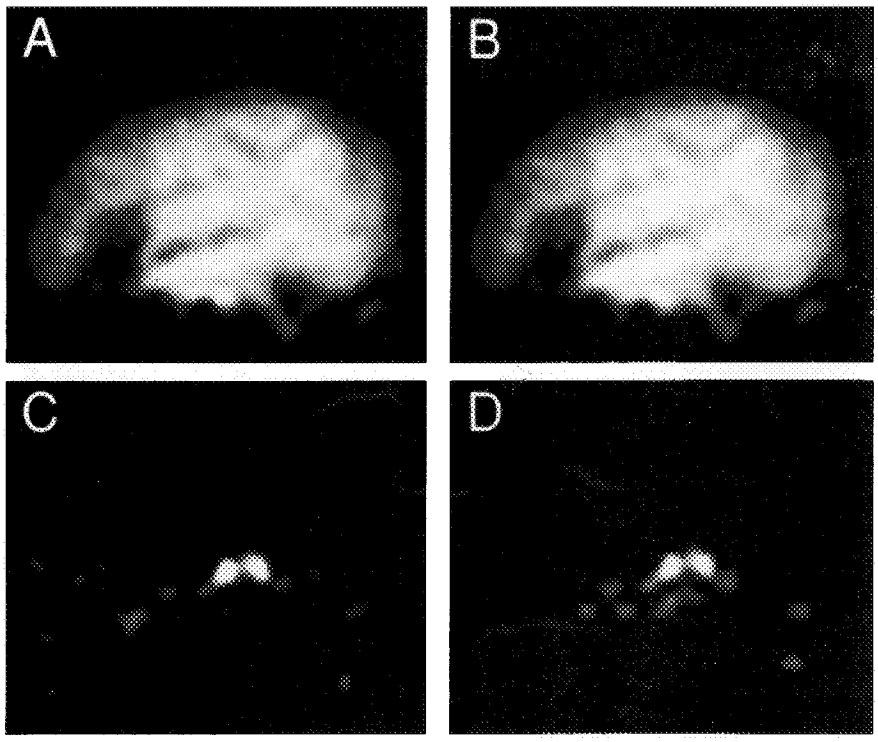
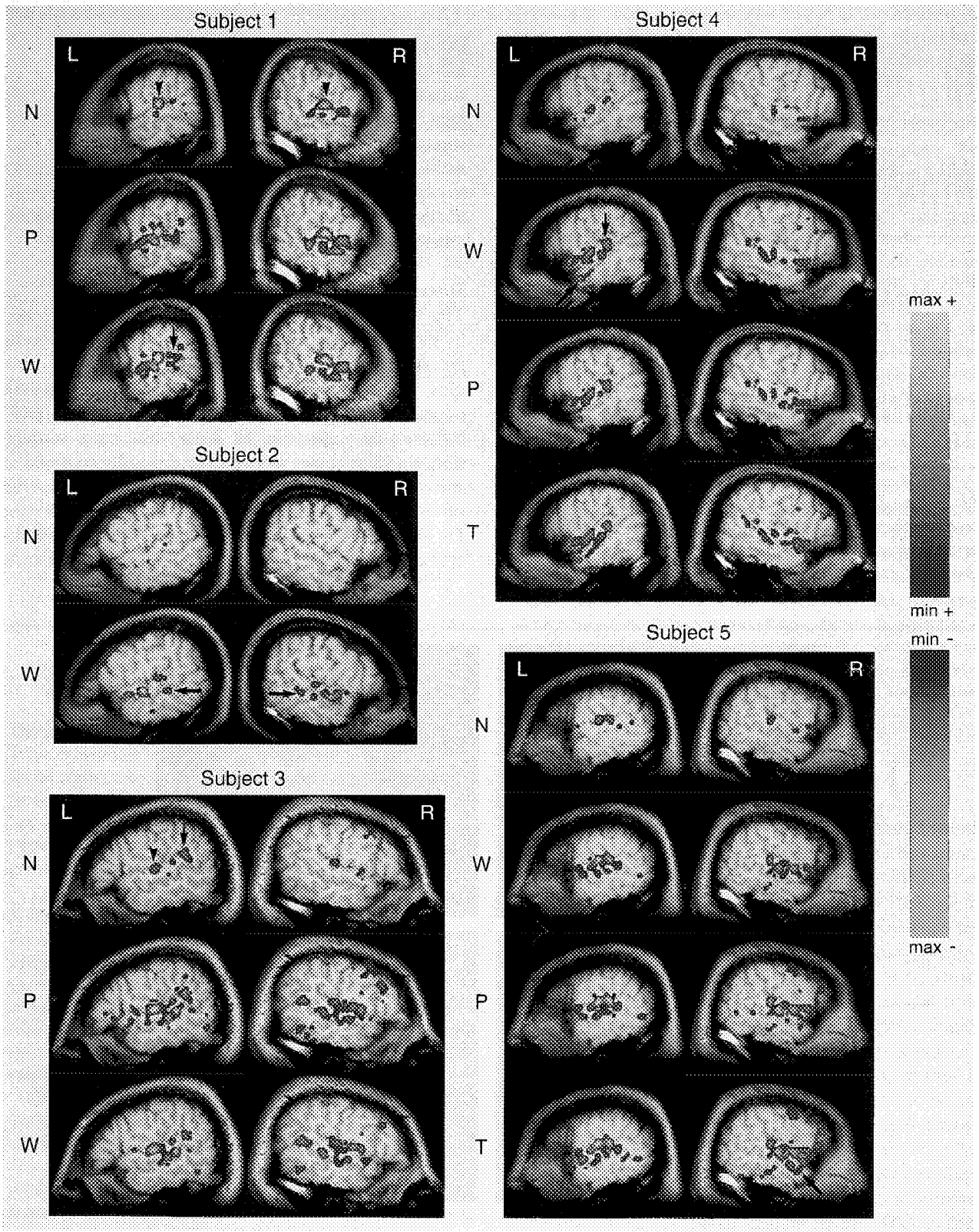


Fig 3. Echoplanar image data (Subject 5, pseudowords, left hemisphere) showing two methods of functional image formation. (A) "Baseline image" obtained by averaging 10 images taken during baseline periods. (B) "Activation image" obtained by averaging 10 images taken during peak activation periods. Signal increases of 1 to 5% in the superior temporal gyrus during activation (see Fig 2) are virtually imperceptible. (C) "Subtraction image" obtained by subtracting A from B and excluding differences less than 1%. (D) Functional image created by the cross-correlation technique: pixel brightness values = $K(r\sigma)$ for all pixels with $|r| \geq 0.50$.


Subject 1
L
R
N
P
W
Subject 2
L
R
N
W
Subject 3
L
R
N
P
W
Subject 4
L
R
N
W
P
T
Subject 5
L
R
N
W
P
T
max +
min +
min -
max -

rior or posterior to the TTG. Responses to speech stimuli were seen over a considerably wider area in all subjects, typically including most of the dorsal STG and clearly spreading to the superior temporal sulcus (STS) in all subjects. Although most of the STS activity appeared to be centered on the sulcus or on its superior aspect, several subjects also showed foci of activity in either the posterior or anterior (see Fig 4) middle temporal gyrus. There were no consistent response differences using different speech stimuli. Activation area tended to be larger during pseudoword presentation than during word presentation (Subjects 1, 3, and 5), although this difference was not statistically significant (see Fig 5) and was absent in Subject 4. STS activation tended to be more extensive during stimulation with text than with words or pseudowords (see Fig 4), although this difference was also small.

Left-Right Asymmetry

Activation in the temporal lobes was approximately symmetrical in all subjects. Subjects 3, 4, and 5 showed small foci (3–8 voxels) in the left STG with activation values exceeding any obtained in the right hemisphere. This phenomenon is best represented using cumulative response value distributions, as shown in Figure 6 for subjects tested with both pseudowords and words. Subject 4 showed the clearest asymmetry, with approximately 20% of voxels on the left side exceeding maximum right hemisphere values.

The total temporal lobe area (number of voxels) activated by speech stimuli was roughly symmetrical in all subjects except Subject 5, who showed a greater number of active voxels on the left side (mean for the three speech conditions = 36) than on the right (mean = 26). In several subjects there was a focal asymmetry consisting of more extensive activation posterior to TTG in the left hemisphere (Subjects 3, 4, and 5).

Fig 4. Brain activation images, Subjects 1–5 (see text for details). Positively correlated pixels are displayed in red-yellow, *negatively correlated pixels in* blue-cyan, *with colors reflecting magnitude and sign of* (rσ). *The same scale applies to all subjects. The* arrowheads *indicate the transverse temporal gyrus (TTG) in Subjects 1 and 3. An area posterior to the TTG appears larger on the left side in Subjects 1, 3, and 4* (vertical arrows). *Small activation foci in the posterior middle temporal gyrus are visible in several subjects* (horizontal arrows, *Subject 2). Small activation foci in the anterior middle temporal gyrus are visible in Subjects 4 and 5* (diagonal arrows). *Right lateral frontal activity is apparent in Subjects 3, 4, and 5. A linear group of negatively correlated pixels in the right hemisphere of Subject 5 is noted, possibly representing a superficial cortical vein. The* small black squares *in one left brain image of Subject 5 (pseudoword condition) identify corner locations of the 16-voxel region represented in Figure 2.* L = *left;* R = *right;* N = *noise;* P = *pseudowords;* W = *words.*

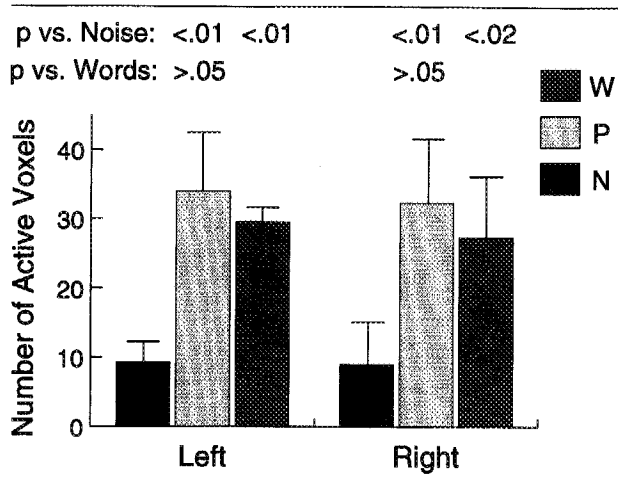


Fig 5. Mean number of active temporal lobe pixels during three stimulus conditions in Subjects 1, 3, 4, and 5. Both word and pseudoword conditions differed significantly from noise bilaterally, while no significant differences were seen between word and pseudoword conditions. N = *noise;* P = *pseudowords;* W = *words.*

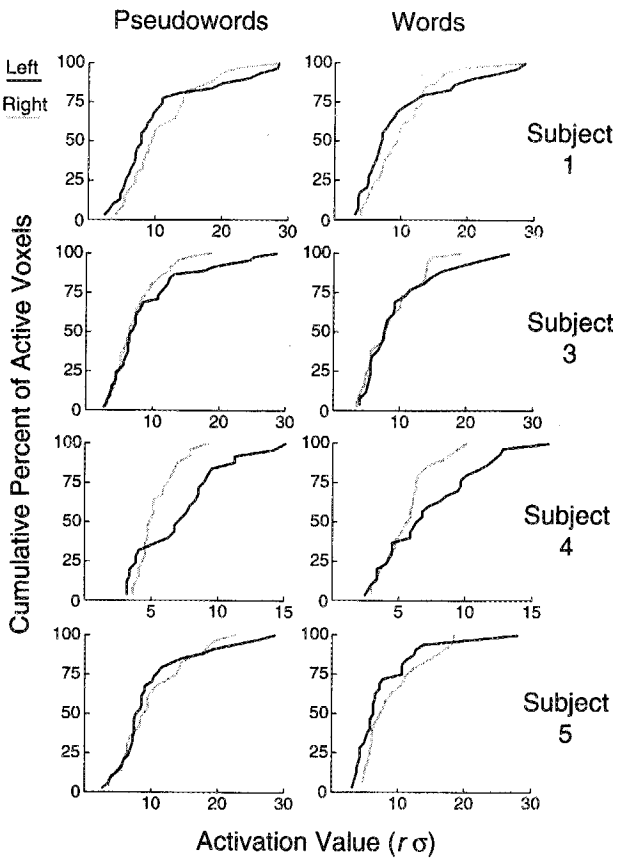


Fig 6. Cumulative percent distributions of (r σ) *values in the right and left temporal lobes during two speech conditions. Peak values observed in the left hemisphere exceeded right hemisphere values in Subjects 3, 4, and 5.*

Four subjects showed varying degrees of right frontal lobe activation, particularly in response to speech stimuli, which was not present on the left side (see Fig 4). Most of this activity was scattered about the precentral sulcus, in an area approximating the inferior aspect of Brodmann area 6 in the middle frontal gyrus. The magnitude of this right frontal response varied from trial to trial and roughly paralleled the degree of response in both temporal lobes (e.g., Subjects 3 and 5).

Intersubject Variability and Other Results

Figure 4 demonstrates the considerable topographical variation in activity among subjects. Specific examples include differences in the extent of activity along the STS (compare Subject 1, left, to Subject 3) and differences in the extent of activity in the anterior STG (compare Subject 3 to Subjects 4 and 5). Total area of activity was considerably smaller in Subject 2 than in other subjects, while the overall magnitude of response was considerably less in Subject 4 than in others (see Fig 6).

Signal intensity in some voxels diminished during stimulus presentation, a phenomenon noted previously [36]. These voxels were typically solitary and located next to voxels showing strong positive activation (Subjects 3 and 4). In the right hemisphere of Subject 5, a linear group of these negatively correlated voxels was observed, having a location and orientation corresponding to the inferior anastomotic vein of Labbé.

Discussion

Our results demonstrate the feasibility of using FMRI to investigate the functional anatomy of human audition. Despite unusual conditions including continuous background scanner noise, the method detected activity in both primary and association auditory cortex in response to a wide range of speech and nonspeech stimuli. The method controlled for the effects of scanner noise by treating this stimulus as a constant throughout all baseline and activation periods. The periodic signal changes measured may therefore be attributed to *additional* brain state changes associated with introduction of the experimental stimuli against a constant background of scanner noise. That these responses varied significantly with different types of experimental stimuli suggests that they depend on specifiable stimulus properties, and are not nonspecific responses.

The results reported here corroborate existing knowledge from positron emission tomography (PET) experiments using related stimuli [11–14, 22]. A major finding in the present study and in prior PET studies is that speech stimuli activate a spatially more extensive region of the temporal lobe than does unmodulated noise, and different speech stimuli (words, nonwords, text) produce similar or identical activations [12, 14, 22]. While consistent, these results are not entirely expected given the differences in semantic content that presumably distinguish nonwords from words. Unlike meaningless stimuli, words have learned associations with other symbols, concepts, or responses; these associations, and the processes by which they are activated by phoneme or printed letter clusters, may be described as a *semantic system* [39–43]. Previous studies revealed a frontal region related to this semantic system using word fluency tasks, which require the activation of previously learned associations [13]. While most apparent during tasks requiring overt responses to meaningful stimuli, the semantic system may also operate relatively automatically during passive stimulus exposure, influencing subsequent brain events without subjects' awareness of any influence [44, 45]. That this semantic system should remain unseen after comparing activation by words and nonwords has at least two possible interpretations.

One interpretation is that words and nonwords actually differ little in semantic content, both activating a semantic system located primarily in the STG of one or both hemispheres [14]. Hierarchical parallel processing models of word perception allow for the activation of higher-order codes (such as whole-word codes) by nonword input [46], and there is accumulating evidence for common mechanisms governing recognition of printed words and nonwords [47]. As a conscious correlate of the activation of higher-order codes by nonwords, subjects occasionally report being "reminded of" a real word after hearing a nonword. Although persuasive, these lines of evidence must be reconciled with lesion studies showing that large injuries confined to the STG of the left or both hemispheres produce little, if any, difficulty related to semantic association [5, 6]. Such lesions result in the classic syndrome of "pure word deafness," whereas semantic deficits typical of Wernicke aphasia require much larger lesions of the temporal or inferior parietal lobes beyond the STG [48–51], and may result from inferior temporal lobe lesions quite distant from the STG [52, 53].

A second interpretation is that the stimuli used here did not succeed in activating the semantic system to a level detectable by the FMRI technique. The semantic system may be a relatively distributed neural network covering a fairly large cortical region [48, 53, 54]; blood flow changes within such a network may not be as focally concentrated as is the case for unimodal cortex, making them less likely to be detected. In support of this view, PET studies have typically shown much lower levels of blood flow change during purely cognitive tasks than during sensory or motor tasks [13]. This

interpretation would help reconcile data from lesion studies, which suggest major participation of the posterior dominant temporal and parietal lobes in language, with data from PET studies, which generally show little activity in these same areas during semantic tasks. Recent PET activations using more complex tasks demonstrated semantic task-specific activation in temporoparietal and ventral temporal foci of the left hemisphere, suggesting that the posterior semantic networks indicated by lesion studies may also be detectable under certain circumstances using blood flow measurements [54].

Differences between responses to speech and nonspeech stimuli could be due to variables unrelated to semantic processing. In acoustic terms, these types of stimuli differ markedly in the amount of frequency modulation present in each, since speech is characterized by continuous rapid transitions in the spectral position of overtone clusters (formants) present in the auditory signal [55]. These temporal acoustic features are the physical equivalents of perceived consonant and vowel sounds, and as such must be transcoded for phoneme perception to occur [56]. The main site of this acoustic analysis is likely to be in the STG, where single-unit studies in animals show neurons tuned to specific frequency transitions like those found in speech [18–21], where evoked potentials can be recorded in humans in response to frequency-modulated stimuli [9, 57], and where lesions produce deficits of phoneme discrimination [3–6]. The results of the present study, showing significantly greater activation in the STG by speech than by unmodulated noise stimuli, are consistent with data from these other sources, and support the hypothesis that a major function of unimodal auditory association cortex (Brodmann area 22) is the analysis of temporal acoustic features of speech and other highly modulated sounds.

Test order is a potential confounding factor that may have produced apparent differences between stimulus conditions. PET studies of motor cortex have shown that a decrement in the extent of functional activation occurs with repeated exposure to the same activation procedure [58], possibly due to habituation factors or synaptic plasticity. We controlled for this effect in the word/pseudoword comparison by varying the order of these conditions across subjects. Similarly, the white noise condition was positioned first in the order as this stimulus was expected to show the smallest activation. The smaller activation seen in the white noise condition is therefore unlikely to have resulted from habituation or other nonspecific effects of test order.

Functional imaging studies have often shown asymmetrical activation during stimulation with speech [10, 11, 13, 14, 22]. The present data, suggesting asymmetries favoring the left side in individual cases, are consistent with this previous work. The finding of higher peaks of activity on the left in several subjects could be explained by greater blood flow changes in these foci, by a greater degree of contiguity of activated areas causing less intravoxel averaging of the signal [28], or by combinations of these factors.

Right lateral frontal activity observed in most of our subjects may reflect mechanisms mediating general arousal and attention. Right frontal lesions are believed to result in specific deficits of arousal and poor performance on vigilance tasks [59–61]. Blood flow studies have shown activation in a similar area during tasks specifically probing sustained attention [62–64]. Differences in the level of arousal from trial to trial in the present study may have been responsible for minor differences observed across speech conditions in some individuals. In particular, it seems likely that the pseudoword stimuli may have induced an alerting response, explaining the slightly greater response to pseudowords in several subjects.

The cause of signal decrements observed in some voxels during activation periods is unclear [36]. When these are located immediately adjacent to strongly activated voxels, one explanation is a "steal" phenomenon resulting from local shunting of oxygenated blood toward active voxels and away from neighboring inactive voxels. We observed in 1 subject a different pattern of negatively correlated voxels, consisting of a linear array not adjacent to positive voxels (see Fig 4). This structure could represent a cortical vein draining the active region and subject to increased blood flow during activation periods. The observed signal decrement might therefore be due to "time-of-flight," dephasing, or other flow-related effects. Investigation of these and other hypotheses awaits further study.

As in previous FMRI studies of motor and visual areas, the activated regions in our study appeared to involve cortical tissue almost exclusively [25–29]. Activity was often concentrated along sulci, giving a somewhat "patchy" rather than a homogeneous appearance to the active areas. This appearance likely reflects the fact that active cortical tissue crosses, to some degree, in or out of the slice plane, depending on the configuration of the gyral surface, the angle of sulci with respect to the slice orientation, the variation in depth along the length of sulci, and other morphological factors. The tomographic method used in this study is therefore limited in its ability to fully represent homogeneous, contiguous regions of activation on the convoluted cortical surface. This limitation, while contributing to the inhomogeneous appearance of the activity distributions in Figure 4, should not affect the ability to qualitatively localize activity or to measure large differences in activity distributions between stimulus conditions.

Several other methodological limitations of the study require discussion. First, computer limitations restricted scanning to two brain slices; other potentially active brain regions were therefore not evaluated. In some subjects other regions of interest were imaged during separate stimulus presentations; these results will be the subject of subsequent reports. Instrumentation improvements at our facility recently included development of whole-brain imaging, which should eliminate slice limitations in future studies. Second, no formal auditory acuity testing of subjects was performed. This deficiency seems unlikely to have influenced results as no subject had any symptom related to hearing loss, and all reported hearing the stimuli clearly. Third, as mentioned previously, attention and arousal variables were not controlled. This was a necessary consequence of the passive listening paradigm, which was thought to be most suitable for a "first look" at auditory activation using FMRI. Finally, precise brain slice position varied somewhat across and within subjects, despite efforts to control this using standard measurements from the lateral brain surface. This outcome is a further indication of individual variability in brain shape and should be overcome in future studies using whole-brain acquisitions.

FMRI is an important new noninvasive tool for investigating structure-function relationships in the human brain. FMRI provides high-resolution functional images that are readily coregistered with anatomical images, and is apparently capable of far greater spatial resolution than that reported here [28]. The method is safe and well tolerated, allowing considerable freedom of repeated testing in individuals. FMRI does not require pooling of responses across subjects, enabling detailed study of individual variations in functional anatomy and avoiding potential false-negative results arising from averaging on standardized templates [65]. The FMRI technique should eventually enjoy widespread use, given its relatively low cost and the possibility of its implementation on commercial scanners such as the one used in this study. This availability, together with the other capabilities mentioned, could lead to significantly improved understanding of extremely complex human brain functions such as speech perception.

This work was supported by a grant from the McDonnell-Pew Program in Cognitive Neuroscience and by grant CA41464 from the National Institutes of Health.

Appendix: Text Sentences

Wearing his black rubber boots and parka, Tom started through the farmyard, crossing a stretch of ice first, then sticky mud.

As he approached the cattle yard, Tom saw five or six cows contentedly chewing at bales of hay, their smooth brown coats warmed by the sun.

The dairy maids, splashing through the mud with their naked white legs not yet sunburned, drove the mooing calves out into the great field.

Tom's big horse sank up to its fetlocks in the mud, and as each foot was pulled out of the half-frozen ground, it made a wet, smacking noise.

Riding through the woods, he was delighted by the sight of the old trees, with the moss reviving on their bark and their green buds swelling up to bloom.

To avoid trampling the clover, Tom rode through the shallow streams of the forest, frightening two ducks as his horse mounted the slippery bank.

On the hill, he could already see the meadow far below, part of it in shadow, where the laborers were just beginning their day's mowing.

Dismounting and tethering his horse, Tom walked off into the vast gray-green sea of the meadow, the silky grass reaching as high as his waist.

As he approached he saw more and more peasants working in a long row, each one swinging his scythe in his own way, clearing the grass in wide swaths.

Tom took up a scythe himself and began mowing, as the peasants came out onto the road, sweaty and cheerful, and laughingly greeted him.

References

1. Braak H. On magnopyramidal temporal fields in the human brain—probable morphological counterparts of Wernicke's sensory speech region. Anat Embryol (Berl) 1978;152:141–169
2. Galaburda A, Sanides F. Cytoarchitectonic organization of the human auditory cortex. J Comp Neurol 1980;190:597–610
3. Kanshepolsky J, Kelley JJ, Waggener JD. A cortical auditory disorder: clinical, audiologic and pathologic aspects. Neurology 1973;23:699–705
4. Goldstein M. Auditory agnosia for speech ("pure word deafness"): a historical review with current implications. Brain Lang 1974;1:195–204
5. Auerbach SH, Allard T, Naeser M, et al. Pure word deafness. Analysis of a case with bilateral lesions and a defect at the prephonemic level. Brain 1982;105:271–300
6. Tanaka Y, Yamadori A, Mori E. Pure word deafness following bilateral lesions: a psychophysical analysis. Brain 1987;110:381–403
7. Celesia GG. Organization of auditory cortical areas in man. Brain 1976;99:403–414
8. Romani GL, Williamson SJ, Kaufman L, Brenner D. Characterization of the human auditory cortex by the neuromagnetic method. Exp Brain Res 1982;47:381–393
9. Hari R. Activation of human auditory cortex by speech sounds. Acta Otolaryngol Suppl (Stockh) 1991;491:132–138
10. Greenberg JH, Reivich M, Alavi A, et al. Metabolic mapping of functional activity in human subjects with the (^{18}F)-fluorodeoxyglucose technique. Science 1981;212:678–680
11. Mazziotta JC, Phelps ME, Carson RE, Kuhl DE. Tomographic mapping of human cerebral metabolism: auditory stimulation. Neurology 1982;32:921–937
12. Lauter J, Herscovitch P, Formby C, Raichle ME. Tonotopic

organization of human auditory cortex revealed by positron emission tomography. Hear Res 1985;20:199–205
13. Petersen SE, Fox PT, Posner MI, et al. Positron emission tomographic studies of the cortical anatomy of single-word processing. Nature 1988;331:585–589
14. Wise R, Chollet F, Hadar U, et al. Distribution of cortical neural networks involved in word comprehension and word retrieval. Brain 1991;114:1803–1817
15. Pandya DN, Sanides F. Architectonic parcellation of the temporal operculum in rhesus monkey and its projection pattern. Z Anat Entwicklungsgesch 1973;139:123–161
16. Galaburda AM, Pandya DN. The intrinsic architectonic and connectional organization of the superior temporal region of the rhesus monkey. J Comp Neurol 1983;221:169–184
17. Morel A, Kaas JH. Subdivisions and connections of auditory cortex in owl monkeys. J Comp Neurol 1992;318:27–63
18. Whitfield IC, Evans EF. Responses of auditory cortical neurons to stimuli of changing frequency. J Neurophysiol 1965;28: 655–672
19. Winter P, Funkenstein HH. The effect of species-specific vocalization on the discharge of auditory cortical cells in the awake squirrel monkey (Saimiri sciureus). Exp Brain Res 1973;18: 489–504
20. Mendelson JR, Cynader MS. Sensitivity of cat primary auditory cortex (Al) to the direction and rate of frequency modulation. Brain Res 1985;327:331–335
21. Steinschneider M, Arezzo JC, Vaughan HG. Tonotopic features of speech-evoked activity in primate auditory cortex. Brain Res 1990;519:158–168
22. Zatorre RJ, Evans AC, Meyer E, Gjedde A. Lateralization of phonetic and pitch discrimination in speech processing. Science 1992;256:846–849
23. Fox PT, Raichle ME. Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. Proc Natl Acad Sci USA 1986; 83:1140–1144
24. Ogawa S, Lee TM, Nayak AS, Glynn P. Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. Magn Reson Med 1990;14:68–78
25. Bandettini PA, Wong EC, Hinks RS, et al. Time course EPI of human brain function during task activation. Magn Reson Med 1992;25:390–397
26. Kwong KK, Belliveau JW, Chesler DA, et al. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. Proc Natl Acad Sci USA 1992;89:5675–5679
27. Ogawa S, Tank DW, Menon R, et al. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping using MRI. Proc Natl Acad Sci USA 1992;89:5951–5955
28. Frahm J, Merboldt K-D, Hänicke W. Functional MRI of human brain at high spatial resolution. Magn Reson Med 1993;29:139–144
29. Rao SM, Binder JR, Bandettini PA, et al. Functional magnetic resonance imaging of complex human movements. Neurology 1993;43:2311–2318
30. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 1971;9:97–113
31. Wong EC, Jesmanowicz A, Hyde JS. Coil optimization for MRI by conjugate gradient descent. Magn Reson Med 1991;21: 39–48
32. Wong EC, Boskamp E, Hyde JS. A volume optimized quadrature elliptical endcap birdcage brain coil. In: Book of abstracts, 11th annual meeting, Society for Magnetic Resonance in Medicine (SMRM). Berkeley, CA: SMRM, 1992:4015
33. Jesmanowicz A, Wong EC, DeYoe EA, Hyde JS. Method to correct anatomic distortion in echo planar images. In: Book of abstracts, 11th annual meeting, Society for Magnetic Resonance in Medicine (SMRM). Berkeley, CA: SMRM, 1992:4260
34. Francis WN, Kucera H. Frequency analysis of English usage: lexicon and grammar. Boston: Houghton Mifflin, 1982
35. DeYoe EA, Neitz J, Bandettini PA, et al. Time course of event-related MR signal enhancement in visual and motor cortex. In: Book of abstracts, 11th annual meeting, Society for Magnetic Resonance in Medicine (SMRM). Berkeley, CA: SMRM, 1992: 1824
36. Menon RS, Ogawa S, Kim S-G, et al. Functional brain mapping using magnetic resonance imaging. Signal changes accompanying visual stimulation. Invest Radiol 1992;27:S47–S53
37. Bandettini PA, Jesmanowicz A, Wong EC, Hyde JS. Processing strategies for time-course data sets in functional MRI of the human brain. Magn Reson Med 1993;30:161–173
38. Binder JR, Jesmanowicz A, Rao SM, et al. Analysis of phase differences in periodic functional MRI activation data. In: Proceedings of the 12th annual meeting, Society for Magnetic Resonance in Medicine (SMRM). Berkeley, CA: SMRM, 1993:1383
39. Sidman M. Reading and auditory-visual equivalences. J Speech Hear Res 1971;14:5–13
40. Tulving E. Episodic and semantic memory. In: Tulving E, Donaldson W, eds. Organization of memory. New York: Academic, 1972:381–403
41. Rosch E. Cognitive representation of semantic categories. J Exp Psychol [Gen] 1975;104:192–233
42. Warrington EK. Neuropsychological studies of verbal semantic systems. Philos Trans R Soc London [Biol] 1981; 295:411–423
43. Shallice T. Specialisation within the semantic system. Cogn Neuropsychol 1988;5:133–142
44. Carr TH, McCauley C, Sperber RD, Parmalee CM. Words, pictures, and priming: on semantic activation, conscious identification, and the automaticity of information processing. J Exp Psychol [Human Percept] 1982;8:757–777
45. Marcel AJ. Conscious and unconscious perception: experiments on visual masking and word recognition. Cogn Psychol 1983;15: 197–237
46. McClelland JL, Rumelhart DE. An interactive activation model of context effects in letter perception: part 1. An account of basic findings. Psychol Rev 1981;88:375–407
47. Humphreys GW, Evett LJ. Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual-route theory of reading. Behav Brain Sci 1985;8:689–740
48. Geschwind N, Quadfasel FA, Segarra JM. Isolation of the speech area. Neuropsychologia 1968;6:327–340
49. Naeser M, Hayward RW, Laughlin SA, Zatz LM. Quantitative CT scan studies of aphasia. I. Infarct size and CT numbers. Brain Lang 1981;12:140–164
50. Selnes OA, Niccum N, Knopman DS, Rubens AB. Recovery of single word comprehension: CT-scan correlates. Brain Lang 1984;21:72–84
51. Hart J, Gordon B. Delineation of single-word semantic comprehension deficits in aphasia, with anatomic correlation. Ann Neurol 1990;27:226–231
52. Kertesz A, Sheppard A, MacKenzie R. Localization in transcortical sensory aphasia. Arch Neurol 1982;39:475–478
53. Alexander MP, Hiltbrunner B, Fischer RS. Distributed anatomy of transcortical sensory aphasia. Arch Neurol 1989;46:885–892
54. Démonet J-F, Chollet F, Ramsay S, et al. The anatomy of phonological and semantic processing in normal subjects. Brain 1992; 115:1753–1768
55. Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M. Perception of the speech code. Psychol Rev 1967;74:431–461
56. Kay RH. Hearing of modulation in sounds. Physiol Rev 1982; 62:894–975
57. Stefanatos GA, Green GGR, Ratcliff GG. Neurophysiological evidence of auditory channel anomalies in developmental dysphasia. Arch Neurol 1989;46:871–875

58. Seitz RJ, Roland PE, Bohm C, et al. Motor learning in man: a positron emission tomographic study. Neuroreport 1990;1: 57–66
59. Heilman KM, Watson RT, Valenstein E. Neglect and related disorders. In: Heilman KM, Valenstein E, eds. Clinical neuropsychology. New York: Oxford, 1985;243–293
60. Coslett HB, Bowers D, Heilman KM. Reduction in cerebral activation after right hemisphere stroke. Neurology 1987;37: 957–962
61. Wilkins AJ, Shallice T, McCarthy R. Frontal lesions and sustained attention. Neuropsychology 1987;25:359–366
62. Cohen RM, Semple WE, Gross M, et al. Functional localization of sustained attention. Neuropsychiatry Neuropsychol Behav Neurol 1988;1:3–20
63. Deutsch G, Papanicolaou AC, Bourbon T, Eisenberg HM. Cerebral blood flow evidence of right cerebral activation in attention demanding tasks. Int J Neurosci 1988;36:23–28
64. Pardo JV, Fox PT, Raichle ME. Localization of a human system for sustained attention by positron emission tomography. Nature 1991;349:61–64
65. Steinmetz H, Seitz RJ. Functional anatomy of language processing; neuroimaging and the problem of individual variability. Neuropsychologia 1991;29:1149–1161